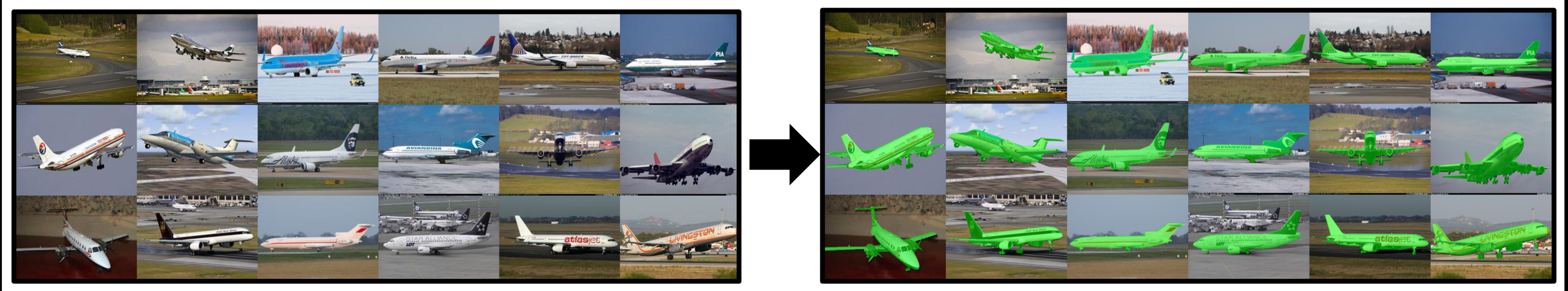
Active Image Segmentation Propagation

Suyog Dutt Jain and Kristen Grauman
University of Texas at Austin
http://vision.cs.utexas.edu/projects/activeseg/

THE UNIVERSITY OF TEXAS AT AUSTIN

Introduction

Weakly Supervised Image Collection → **Goal: Segment Common Object**



A stage-wise algorithm for active human annotation and segmentation propagation in image collections.

Existing methods are either passive or only select annotations during initialization.

Problem

Only weak supervision

Inexpensive but too inaccurate

Human labels everything

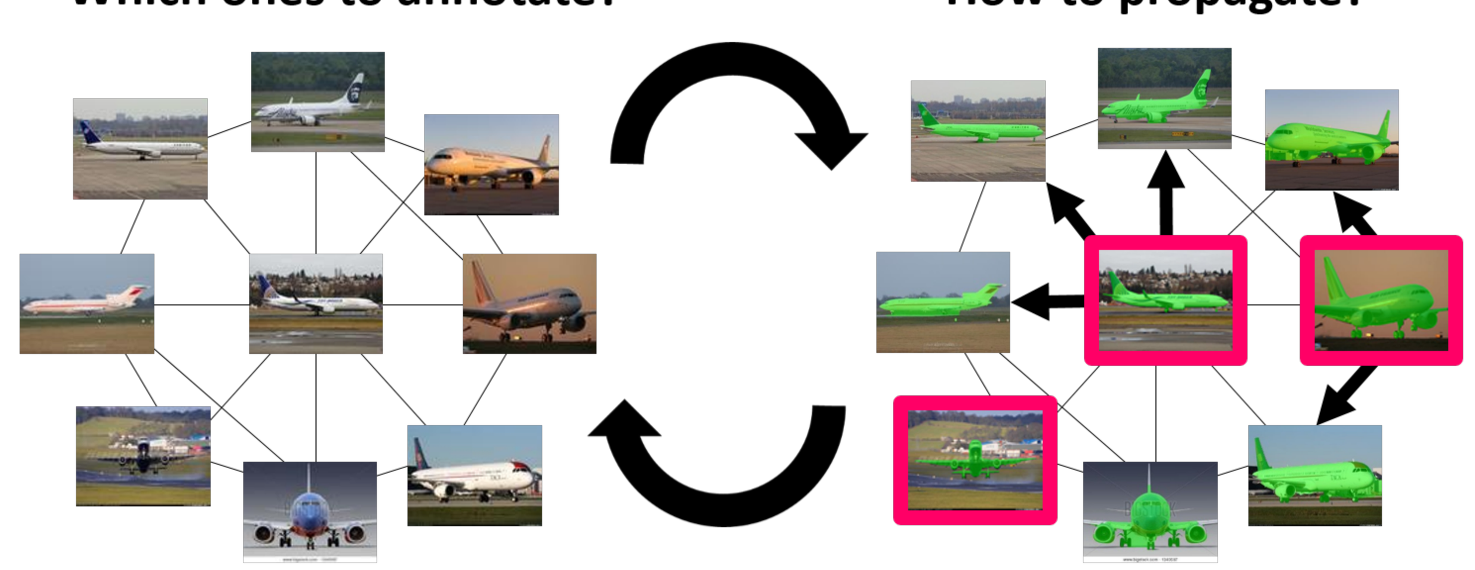
Accurate but too expensive

Microsoft COCO dataset – 2.5M object instances ~ $400,000K

Our Idea

Active Human Annotation | **Joint Segmentation Propagation**

Which ones to annotate? | How to propagate?



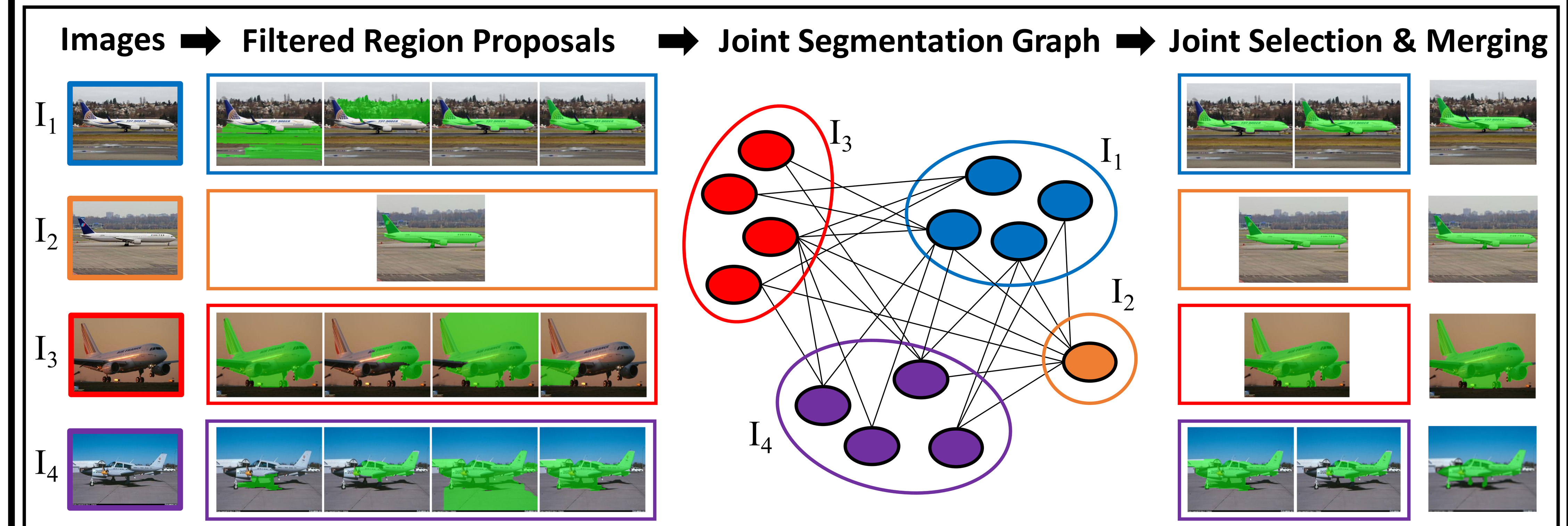
Actively request human annotations for select images | **Update segmentations for unlabeled images**

Our method prioritizes human intervention for those images that are **uncertain** and **influential** in the graph, while also mutually **diverse**.

Through large scale experiments on **nearly 1 million images** we show that **actively allocating human effort** leads to **substantial savings** in annotation costs.

Joint Segmentation Propagation

Markov random field based segmentation propagation using a joint segmentation graph



Unary and Pairwise Potentials

Average region saliency

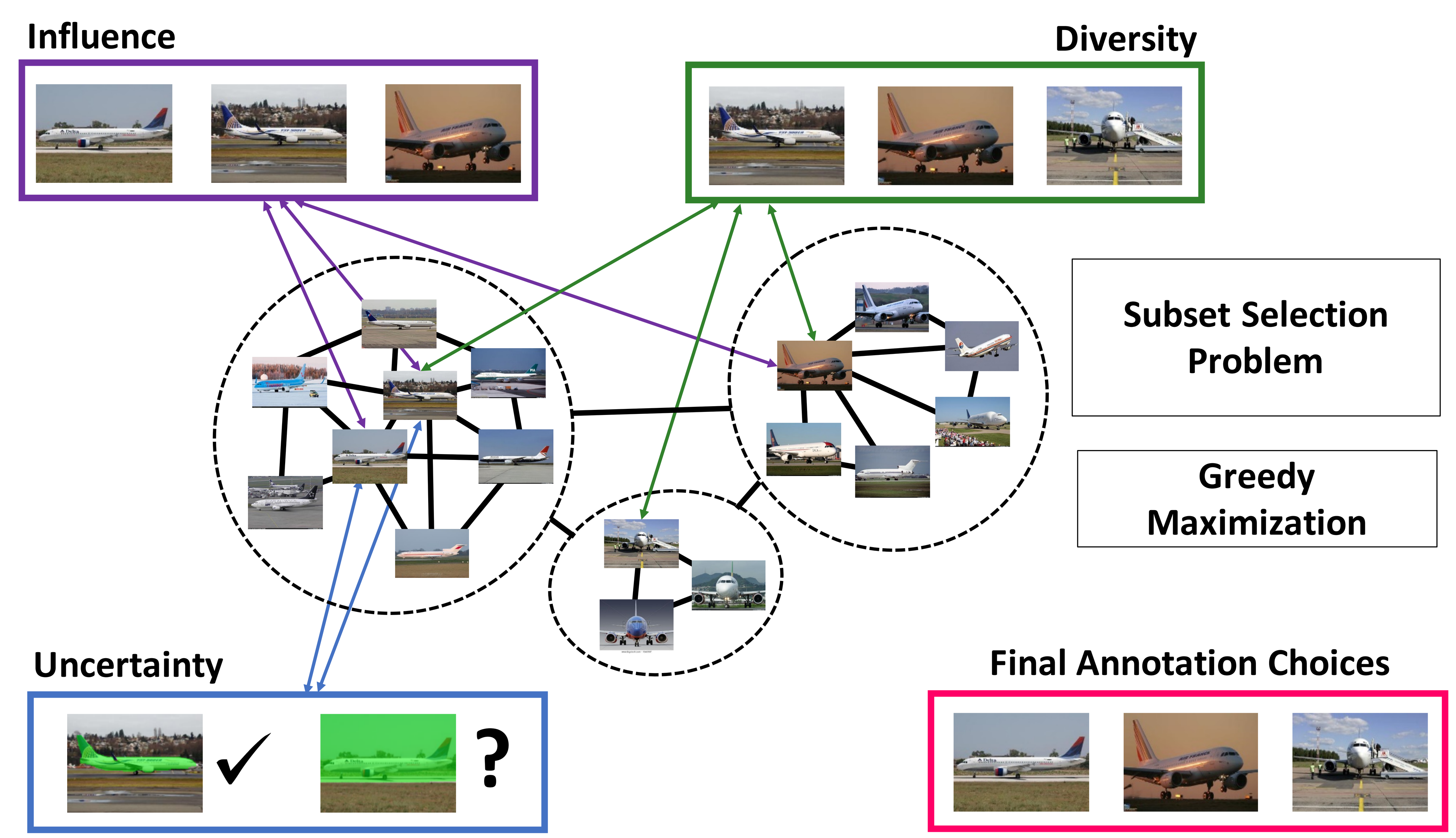
Match scores with human segmented images (using CNN features)

Pairwise region matching scores (using CNN features)

Very efficient (**1 min for 1400 images**) compared to pixel based approach (**225 hours**) [Rubinstein 2012]

Active Human Annotation

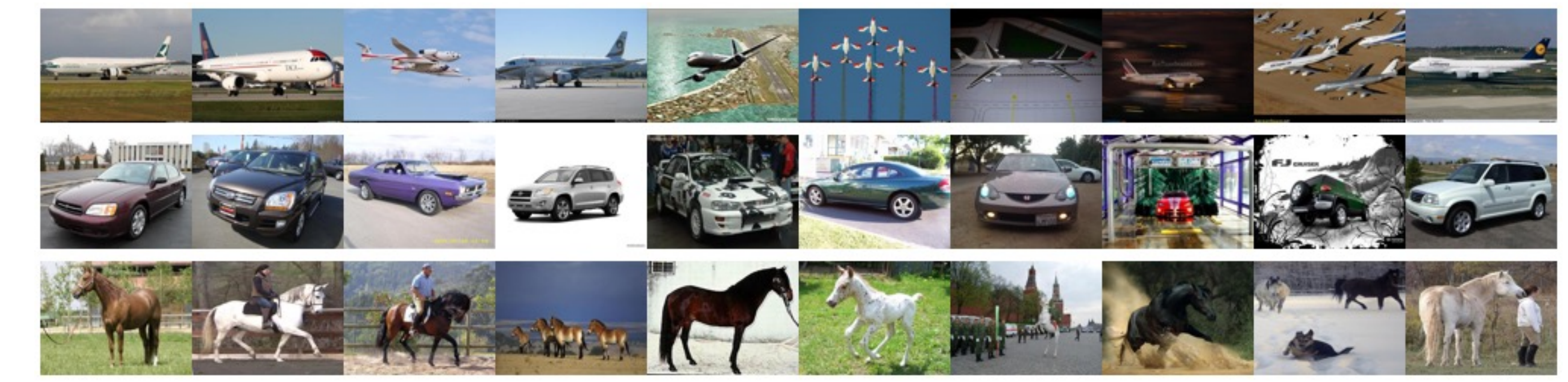
Stage-wise algorithm to actively select images for human annotation



Subset Selection Problem

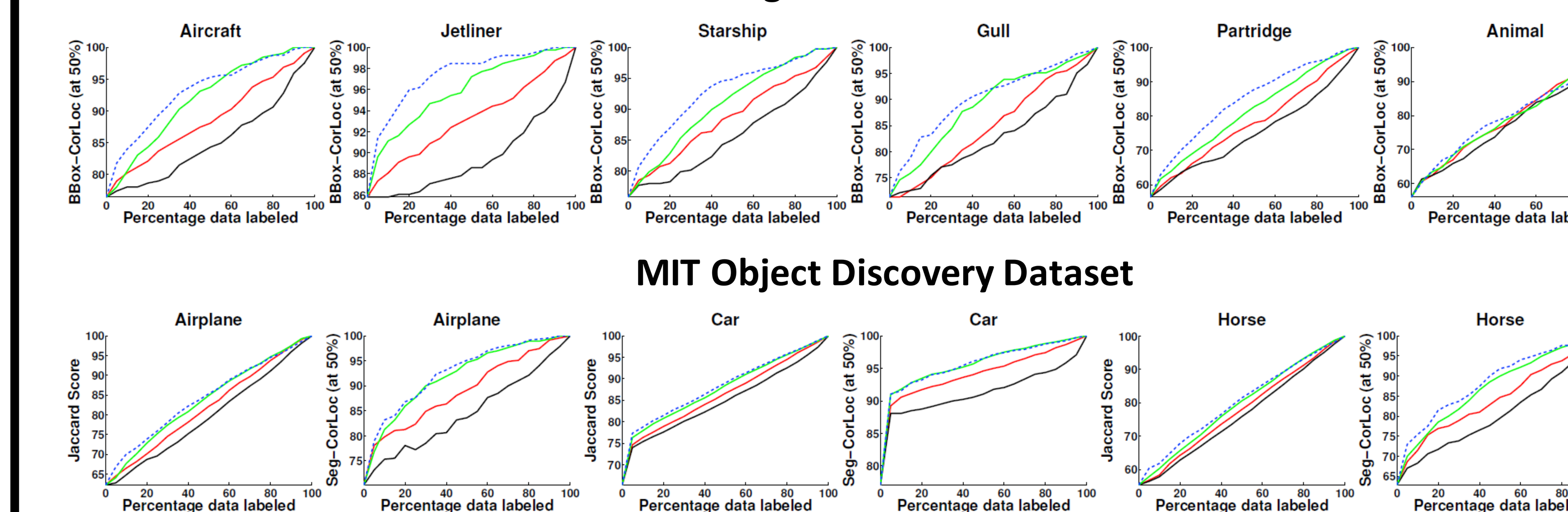
Greedy Maximization

Active Segmentation Propagation - Results



Example choices made by our active annotation algorithm

Quantitative Results

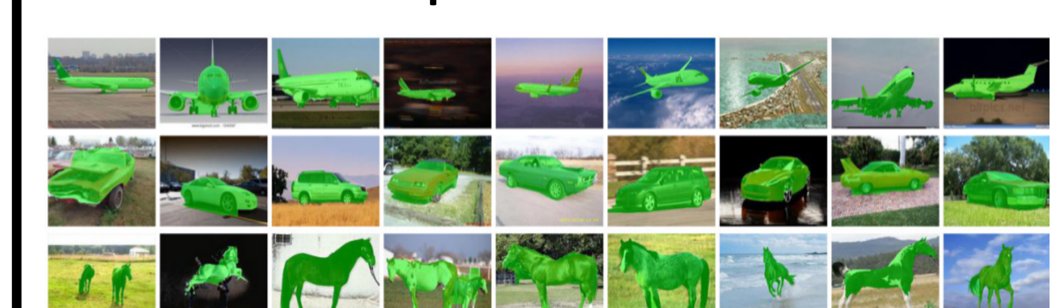


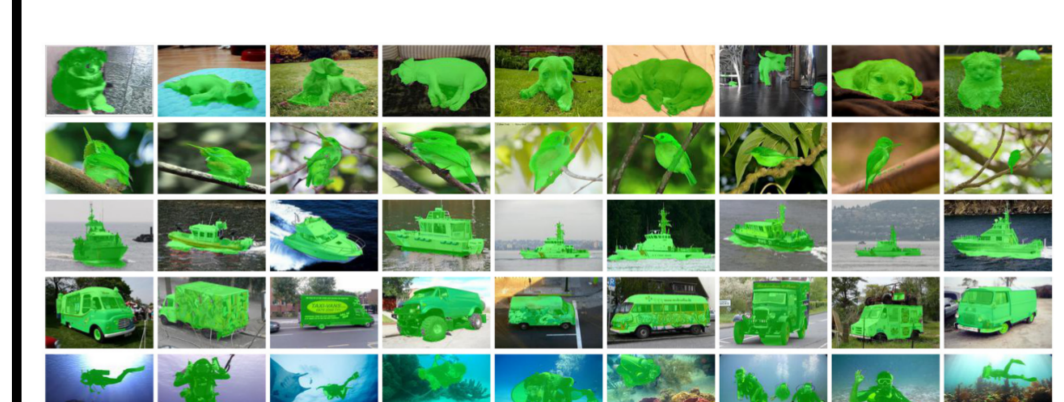
Our proposed active segmentation propagation approach is able to generate high quality segmentations with significantly less human annotation cost.

Outperforms state of the art segmentation propagation approach [ImageNet-SegProp, Guillaumin et al. IJCV 2014] **while requiring 26% less human annotated data.**

Weakly Supervised Segmentation - Results

State of the art performance in most cases





MIT Object Discovery Dataset

Methods	MIT dataset (subset)		
	Airplane	Car	Horse
# Images	82	89	93
Joulin et al. [19]	15.36	37.15	30.16
Joulin et al. [20]	11.72	35.15	29.53
Kim et al. [21]	7.9	0.04	6.43
Rubinstein et al. [35]	55.81	64.42	51.65
Chen et al. [9]	54.62	**69.2**	44.46
Ours	**58.65**	66.47	**53.57**

ImageNet Dataset

# Classes	# Images
3,624	939,516

Methods	BBox-CorLoc
Top obj. box [3]	37.42
Tang et al. [42]	53.20
Ours	**57.64**

With nearly 1 million images, a performance gain of 4.44% means that we correctly localize 41,715 more images.